

9

OIKE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,052

DATE: 07/27/2001

TIME: 11:00:24

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

ENTERED

```

3 <110> APPLICANT: Harada, John
4   Lotan, Tamar
5   Ohto, Masa-Aki
6   Goldberg, Robert B.
7   Fischer, Robert L.
8   Bui, Anhthu
9   Khong, Raymond
10  The Regents of the University of California
12 <120> TITLE OF INVENTION: LEAFY COTYLEDON1 Genes and Their Uses
14 <130> FILE REFERENCE: 023070-077630US
16 <140> CURRENT APPLICATION NUMBER: US 09/516,052
17 <141> CURRENT FILING DATE: 2000-03-01
19 <150> PRIOR APPLICATION NUMBER: US 09/193,931
20 <151> PRIOR FILING DATE: 1998-11-17
22 <150> PRIOR APPLICATION NUMBER: US 09/103,478
23 <151> PRIOR FILING DATE: 1998-06-24
25 <150> PRIOR APPLICATION NUMBER: US 09/026,221
26 <151> PRIOR FILING DATE: 1998-02-19
28 <150> PRIOR APPLICATION NUMBER: US 08/804,534
29 <151> PRIOR FILING DATE: 1997-02-21
31 <160> NUMBER OF SEQ ID NOS: 39
33 <170> SOFTWARE: PatentIn Ver. 2.1
35 <210> SEQ ID NO: 1
36 <211> LENGTH: 627
37 <212> TYPE: DNA
38 <213> ORGANISM: Arabidopsis thaliana
40 <220> FEATURE:
41 <221> NAME/KEY: CDS
42 <222> LOCATION: (1)..(627)
43 <223> OTHER INFORMATION: LEAFY COTYLEDON1 (LEC1)
45 <400> SEQUENCE: 1
46 atg acc agc tca gtc ata gta gcc ggc gcc ggt gac aag aac aat ggt 48
47 Met Thr Ser Ser Val Ile Val Ala Gly Ala Gly Asp Lys Asn Asn Gly
48   1           5           10           15
50 atc gtg gtc cag cag caa cca cca tgt gtg gct cgt gag caa gac caa 96
51 Ile Val Val Gln Gln Gln Pro Pro Cys Val Ala Arg Glu Gln Asp Gln
52           20           25           30
54 tac atg cca atc gca aac gtc ata aga atc atg cgt aaa acc tta ccg 144
55 Tyr Met Pro Ile Ala Asn Val Ile Arg Ile Met Arg Lys Thr Leu Pro
56           35           40           45
58 tct cac gcc aaa atc tct gac gac gcc aaa gaa acg att caa gaa tgt 192
59 Ser His Ala Lys Ile Ser Asp Asp Ala Lys Glu Thr Ile Gln Glu Cys
60           50           55           60
62 gtc tcc gag tac atc agc ttc gtg acc ggt gaa gcc aac gag cgt tgc 240
63 Val Ser Glu Tyr Ile Ser Phe Val Thr Gly Glu Ala Asn Glu Arg Cys
64   65           70           75           80
66 caa cgt gag caa cgt aag acc ata act gct gaa gat atc ctt tgg gct 288

```

RAW SEQUENCE LISTING

DATE: 07/27/2001

PATENT APPLICATION: US/09/516,052

TIME: 11:00:24

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

```

67 Gln Arg Glu Gln Arg Lys Thr Ile Thr Ala Glu Asp Ile Leu Trp Ala
68                85                90                95
70 atg agc aag ctt ggg ttc gat aac tac gtg gac ccc ctc acc gtg ttc 336
71 Met Ser Lys Leu Gly Phe Asp Asn Tyr Val Asp Pro Leu Thr Val Phe
72                100                105                110
74 att aac cgg tac cgt gag ata gag acc gat cgt ggt tct gca ctt aga 384
75 Ile Asn Arg Tyr Arg Glu Ile Glu Thr Asp Arg Gly Ser Ala Leu Arg
76                115                120                125
78 ggt gag cca ccg tcg ttg aga caa acc tat gga gga aat ggt att ggg 432
79 Gly Glu Pro Pro Ser Leu Arg Gln Thr Tyr Gly Gly Asn Gly Ile Gly
80                130                135                140
82 ttt cac ggc cca tct cat ggc cta cct cct ccg ggt cct tat ggt tat 480
83 Phe His Gly Pro Ser His Gly Leu Pro Pro Pro Gly Pro Tyr Gly Tyr
84 145                150                155                160
86 ggt atg ttg gac caa tcc atg gtt atg gga ggt ggt cgg tac tac caa 528
87 Gly Met Leu Asp Gln Ser Met Val Met Gly Gly Gly Arg Tyr Tyr Gln
88                165                170                175
90 aac ggg tcg tcg ggt caa gat gaa tcc agt gtt ggt ggt ggc tct tcg 576
91 Asn Gly Ser Ser Gly Gln Asp Glu Ser Ser Val Gly Gly Gly Ser Ser
92                180                185                190
94 tct tcc att aac gga atg ccg gct ttt gac cat tat ggt cag tat aag 624
95 Ser Ser Ile Asn Gly Met Pro Ala Phe Asp His Tyr Gly Gln Tyr Lys
96                195                200                205
98 tga 627
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 208
103 <212> TYPE: PRT
104 <213> ORGANISM: Arabidopsis thaliana
106 <400> SEQUENCE: 2
107 Met Thr Ser Ser Val Ile Val Ala Gly Ala Gly Asp Lys Asn Asn Gly
108 1 5 10 15
109 Ile Val Val Gln Gln Gln Pro Pro Cys Val Ala Arg Glu Gln Asp Gln
110 20 25 30
111 Tyr Met Pro Ile Ala Asn Val Ile Arg Ile Met Arg Lys Thr Leu Pro
112 35 40 45
113 Ser His Ala Lys Ile Ser Asp Asp Ala Lys Glu Thr Ile Gln Glu Cys
114 50 55 60
115 Val Ser Glu Tyr Ile Ser Phe Val Thr Gly Glu Ala Asn Glu Arg Cys
116 65 70 75 80
117 Gln Arg Glu Gln Arg Lys Thr Ile Thr Ala Glu Asp Ile Leu Trp Ala
118 85 90 95
119 Met Ser Lys Leu Gly Phe Asp Asn Tyr Val Asp Pro Leu Thr Val Phe
120 100 105 110
121 Ile Asn Arg Tyr Arg Glu Ile Glu Thr Asp Arg Gly Ser Ala Leu Arg
122 115 120 125
123 Gly Glu Pro Pro Ser Leu Arg Gln Thr Tyr Gly Gly Asn Gly Ile Gly
124 130 135 140
125 Phe His Gly Pro Ser His Gly Leu Pro Pro Pro Gly Pro Tyr Gly Tyr
126 145 150 155 160

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,052

DATE: 07/27/2001

TIME: 11:00:24

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

```

127 Gly Met Leu Asp Gln Ser Met Val Met Gly Gly Gly Arg Tyr Tyr Gln
128                               165                               170                               175
129 Asn Gly Ser Ser Gly Gln Asp Glu Ser Ser Val Gly Gly Gly Ser Ser
130                               180                               185                               190
131 Ser Ser Ile Asn Gly Met Pro Ala Phe Asp His Tyr Gly Gln Tyr Lys
132                               195                               200                               205
135 <210> SEQ ID NO: 3
136 <211> LENGTH: 3395
137 <212> TYPE: DNA
138 <213> ORGANISM: Arabidopsis thaliana
140 <220> FEATURE:
141 <223> OTHER INFORMATION: 3.4 kb BstYI fragment of genomic DNA containing
142     LEC1 gene
144 <220> FEATURE:
145 <221> NAME/KEY: promoter
146 <222> LOCATION: (1)..(1998)
147 <223> OTHER INFORMATION: LEC1 promoter
149 <220> FEATURE:
150 <221> NAME/KEY: modified_base
151 <222> LOCATION: (762)
152 <223> OTHER INFORMATION: n = g, a, c or t
154 <220> FEATURE:
155 <221> NAME/KEY: CDS
156 <222> LOCATION: (1999)..(2625)
157 <223> OTHER INFORMATION: LEAFY COTYLEDON1 (LEC1)
159 <400> SEQUENCE: 3
160 agatccaaaa caggtcatgg actgggcccgt aaactctatc caaaattctt catgtttttc 60
161 catcttttcaa aaatctttat ccaccattcc attactaggg tgttggtttt attttatttg 120
162 ttgattaatt atgtattaga aaatgtaaag caatattcaa ttgtaacatg catcatctaa 180
163 caccaatata ttgtactaac cttttgtaat ttctctataa acatttttaa aggctaattt 240
164 aaataaaaaa tacaataaac gtgataactc actttcgtaa cgcataattt ttcaaataa 300
165 ccaaaattta ccattttaag taagagaatc tttttaaaat taattttcaa ttcatatta 360
166 taagaaacaa agaatttact gaaacctata ttttattaaa ttttaataaa atatatgact 420
167 aaaataacgt cacgtgaatc tttctcagcc gttcgataat cgaatacttt attgactaag 480
168 tattttattta gaaaatttta aacaacactt aatttctaga aacaaagaga gcctcatatg 540
169 tataaaaatc ttcttcttat ctttctttct ttcttaatat tctttatttt tacttaatta 600
170 ctttggtaat ttgtgaaaaa cacaaccaat gagagaagag cagtttgact ggccacatag 660
171 ccaatgagac aagccaatgg gaaagagata tagagacctc gtaagaaccg ctcttttgcc 720
172 atttgtatca tctctctata aaaccactca accatcaacc tntctttgca tgcaacaaat 780
173 cactcaaata attattttat aaagaacaaa aaaaaaaaga cggcagagaa acaatggaac 840
174 gtggagctcc cttctctcac tatcagctac ccaaattccat ctctggtaat ctaagtggct 900
175 atttgtatac agtatatact tgccctccatg tatattttata ttctcgtgaa aaattggaga 960
176 catgctttat gaattttatg agactttgca acaacgaacg agatgctttc tctctagaaa 1020
177 tttaaattta gatttgtgaa ggttttgga atggcccga gaagacgatt ttatatatac 1080
178 atgcatgcaa gagtttgata tgtatattgt ttcacatgg ctgagtcaaa gttttatcca 1140
179 aatatttoca tgggtgtgta ttagttaaac aaatctctcg tatgtgtcat tgaatatacc 1200
180 cgtgcatgta ccaggaatgt ttttgattct aaaaacgttt ttttctttgt tgtaacggtt 1260
181 gagttttttt cttcgtttca aaacgagatt ctcgtttgtc tcttcccttg tctaaaaaca 1320
182 tctacggttc atgtgattca aaaacactaa aaaaatataa actcattttt ttttaatact 1380

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,052

DATE: 07/27/2001

TIME: 11:00:24

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

```

183 taacatttaa actatatata tatatatata tatatatatc ttataactagt cccaagtttt 1440
184 agtgtgaggt ttttttattc aaaatctatc agtacatttt tlygaaaaga actaagtga 1500
185 attttctcca aattttcctt ttactattga ttttttaatt actggatgtc attaacttta 1560
186 atcttttgat tctttcaacg tttaccattg ggaaccttca catgaaataa atgtctactt 1620
187 tattgagtca taccttcgtc aacataaatt aattgatgtt cttctccaaa ttttgagttt 1680
188 ttggtttttc taataatctt aacgaaagct ttttggtata catgtaaaac gtaacggcaa 1740
189 gaatctgaac agtctactca acgggggtcca taagtctaga atgtagacct cacaacttta 1800
190 ctcttatctt attggtccgt aactaagaac gtgtccctct gattctcttg ttttcttcta 1860
191 attaatctgt atcctacaaa ttttaattatc atttctactt caactaatct ttttttattt 1920
192 cctaaagatt tcaatttctc tctgtatttt ctatgaacag aattgaactt ggaccagcac 1980
193 agcaacaacc caacccaat gaccagctca gtcatagtag ccggcgccgg tgacaagaac 2040
194 aatggtatcg tgggtccagca gcaaccacca tgtgtggctc gtgagcaaga ccaatacatg 2100
195 ccaatcgcaa acgtcataag aatcatgctt aaaaccttac cgtctcacgc caaaatctct 2160
196 gacgacgcca aagaaacgat tcaagaatgt gtctccgagt acatcagctt cgtgaccggt 2220
197 gaagccaacg agcgttgcca acgtgagcaa cgtaagacca taactgctga agatatcctt 2280
198 tgggctatga gcaagcttgg gtctgataac tacgtggacc cctcaccgtt gttcattaac 2340
199 cgttaccgtg agatagagac cgatcgtggt tctgcactta gaggtgagcc accgtcgttg 2400
200 agacaaacct atggaggaaa tgggtattggg ttccacggcc catctcatgg cctacctcct 2460
201 ccgggtcctt atggttatgg tatgttggaac caatccatgg ttatgggagg tggtcggta 2520
202 taccaaaacg ggtcgtcggg tcaagatgaa tccagtgttg gtggtggctc ttcgtcttcc 2580
203 attaacggaa tgccggcttt tgaccattat ggtcagtata agtgaagaag gatttattct 2640
204 tcatTTTTat atctattcaa aacatgtgtt tcatagata ttttattttt atgtcttatt 2700
205 aataacattt ctatataatg ttgcttcttt aaggaaaagt gttgtatgtc aatactttat 2760
206 gagaaaactga tttatatatg caaatgattg aatccaaact gttttgtgga ttaaactcta 2820
207 tgcaacatta tatatttaca tgatctaaag gttttgtaat tcaaaagctg tcatagttag 2880
208 aagataacta aacattgtag taaccaagtt taatttactt ttttgagttt acataactaa 2940
209 ccaagccaaa aggttataaa atctaaatc gttgagttgt caaacttctg aagattgcta 3000
210 tctcttttga gttgctttct tttgggtgct tgagtttcat taggctgagc tgactcgttg 3060
211 ctctctagtc tttcatctct gtcttttcca aggattcata acgttggtcg ctctctgttt 3120
212 ctgcctacac ttcttcaagg gatcattact gaggttaaga gttaaagacc tgaaccatgg 3180
213 ttttctgtaa ctggttcaag ttcattctcc ggttattgtg tggttatctt tcggttagat 3240
214 tgaaacccat atgtttgtc tgtttcttct agttccaagt ttaatttccg gttattgttt 3300
215 ggttttttaa aagtttttaa ggtctattct atgtaaagac tattctacgt acgtacattt 3360
216 atcgcaaaat tgaaagatta taaaaaaaat tgaaa 3395
219 <210> SEQ ID NO: 4
220 <211> LENGTH: 7560
221 <212> TYPE: DNA
222 <213> ORGANISM: Arabidopsis thaliana
224 <220> FEATURE:
225 <223> OTHER INFORMATION: 7.4 kb genomic wild-type fragment containing LEC1
226 gene
228 <220> FEATURE:
229 <221> NAME/KEY: modified_base
230 <222> LOCATION: (1)..(7560)
231 <223> OTHER INFORMATION: n = g, a, c or t
233 <220> FEATURE:
234 <221> NAME/KEY: CDS
235 <222> LOCATION: (4427)..(5054)
236 <223> OTHER INFORMATION: LEAFY COTYLEDON1 (LEC1)

```

RAW SEQUENCE LISTING

DATE: 07/27/2001

PATENT APPLICATION: US/09/516,052

TIME: 11:00:24

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

238 <400> SEQUENCE: 4

```

W--> 239 aattnaccct cactaaaggg aacaaaagct gggtagcggg cccccctcg aggtcgacgg 60
      240 tategataag cttgatatac aattcgtggt cattagaccc ataactatat gacgatgtta 120
      241 aagagaaaat aaatcataaa taaaataaga gtccctatca ataaacctaa ttggctaatt 180
      242 tcaacctcaa agagtagtag gaacaggtaa ggtgaagcca aacagctcct ttacagttg 240
      243 gaccactaga gctgatctgg catacaaagt atgcttattg ggtgtcacg gcccatccgc 300
      244 aaaatgtcgt tggttacgaa gcacccacga catagacggg gccacatgtt agaaaagtgt 360
      245 ttcggcgacg aagattgtgt ccacatcatt agacgtctga actgtccacg tgtctatcaa 420
      246 agctggcgct aaacattacg ttttcgtcgt ttgcgcctcc tagttcacac gtgcaacgaa 480
      247 cgcgtgcgac gtatcaaaat tgttaatttt agccatgtat aaagaatata taaaaatta 540
W--> 248 acctcaggaa tatTTTTgtt ttttcaattg aggccataat atacntnccg atngaaaaat 600
W--> 249 tttncancat atcnctaata tcaaaaaatt atgatgttag taaacgtaaa aaattttacac 660
W--> 250 aaaataantt tcacaaaact tannggggaa attggaacaa anaaaagact ggtgagtgat 720
      251 aagcgatgat ggcgggtgaa tcaggtagcc gtccataaac gtggttgatt ttgagcaaac 780
      252 tctatctac tcttcacact attggaaatc caaaatgtc gtcacaccat aataatgtga 840
      253 attttgttat ggaatttgag ggaacagta gatatatgtt tcaaccagtg aaagttacc 900
W--> 254 tcttttgac atattctacga nagtagaaag tagaaacatt cactaaacgt gacaacttta 960
W--> 255 taaattttct ttttgttaact tttctttaga tttattttac anaagagaaa tataaacgtc 1020
      256 atgctaataa aaaatgcatt atttctacc atctagctag aatattgac aagtcttcac 1080
      257 gttttttgtt tatctcttct ctcataggca tgtccacaaa aggtaagt ttactggttc 1140
      258 aaaatattgc atgagtacta ctaagctcgt atagtttgat cttactatca ttgcgatgag 1200
      259 ggttggttagt ttggaagaaa taaggattta tgcaaatggt aatcattatg tctgctattt 1260
      260 aagaagtaaa ttatgatgct tgttgcggtga acatattaaa tttgcgaaaa ataagcaagg 1320
      261 atacacgaga gaagctcaga tttcacgta acgatgttct atctcttctc attgaggaaa 1380
W--> 262 catatggcca tgatatagct aataagccta cgggattgtc ntttcaacgc cgaatctacc 1440
      263 aaactgttcc atctcttatt atatatagtt tggttattta agtaattaga tgcatacaa 1500
      264 tcttttttct tgccagttgt aatgcagata aaaatatatt ggttggtcta aggattgttc 1560
      265 aaacgtgcct ggttacaaat tattatttat atactttcat ctacatgcga tgcgttattt 1620
      266 ataagataaa aactaagatt tttagttaaa tttataaag agcttacgag ctacaattaa 1680
      267 ttagaaatgg ttgctcagaa atcagaatac tatatatgaa aaaagaagtt ggtatacttg 1740
      268 aaaaaagaaa aaactacttg aaaagatggt aaaagatata gaacgagtat atactttact 1800
      269 caagcacgat agaagtttgt atcaaaacat tgcgttccaa accaatgttt gaagatggtc 1860
      270 aaaggtgcta ctcatgatgt ggtgcgaaga agcttacgaa aaattctgca atgagagata 1920
      271 actttatggg ctgcttggtc aatatattga aaatcatggt agacaacacc aaactctcct 1980
      272 ttaccagaag tcatatttcc ttaacctcag aataagtaaa tcttctagtt tattatttga 2040
      273 aagttgagcg tataattgca atgaaacttt taccaattca ccgcctccta actgagttgt 2100
      274 tgtattatcc tatctcttta gctatccttt ccttgcctct gctccacctg catgtggcct 2160
W--> 275 ctttatttat aatctctcta gattctgcta aagatgtntg ttcaaatgg tttatcttta 2220
      276 aggggaagcaa agtgaatgga aacattttaa gaaaaaaaaa acttttagca gatttccatg 2280
      277 agatttcata ctgatgataa ctaaaataat cttatatgct taagattatt ttagttctaa 2340
      278 acttcatttt gaaatgagag gtcattggcc aggaagatt caatattggt tctttgttaa 2400
      279 ttctcgttgg tttgttttta gtatgggcta gatccaaaac aggtcatgga ctgggccgta 2460
      280 aactctatcc aaaattcttc atgtttttcc atctttcaaa aatctttatc caccattcca 2520
      281 ttactagggt gttggtttta ttttatttgt tgatttagaa aatgtaaaagc 2580
      282 aatattcaat tgtaacatgc atcatctaac accaatatct tgtactaacc ttttgaatt 2640
      283 ttctataaaa ctttttaaaa ggcataattta aataaaaaatt acaataaacg tgataactca 2700
      284 ctttcgtaac gcatatttat tcaaatatac caaaatttac ctttttaagt aagagaatct 2760
      285 ttttaaaaatt aattttcaat ttcattaatt aagaaacaaa gaatttactg aaacctatat 2820
      286 tttattaaat ttttaataaaa tatatgacta aaataacgtc acgtgaatct ttctcagccg 2880

```

VERIFICATION SUMMARY

DATE: 07/27/2001

PATENT APPLICATION: US/09/516,052

TIME: 11:00:25

Input Set : A:\-776-3us.app

Output Set: N:\CRF3\07272001\I516052.raw

L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:239 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:248 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:249 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:250 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:255 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:262 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4